



RECEIVED

JUN 26 2002

TECH CENTER 1600/2900

SEQUENCE LISTING

<110> Von Schaewen, Antje

<120> Plant GntI Sequences and the Use Thereof for the Production of Plants Having Reduced or Lacking N-acetyl Glucosaminyl Transferase I (GnTI) Activity

<130> 032266-003

<140> US 09/591,466

<141> 2000-06-09

<150> EP 98/08001

<151> 1998-09-12

<160> 14

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1669

<212> DNA

<213> Solanum tuberosum

<220>

<221> misc_feature

<222> (659)...(667)

<223> function: Asn codon in this context is a potential glycosylation site;
product: N-glycosylation consensus sequence;
phenotype: N-glycans modulate protein properties;

<220>

<221> misc_feature

<222> (659)...(667)

<223> standard_name: N-glycosylation site;
label: pot-CHO;
note: GnTI-coding sequences from animals do not contain this feature.

<220>

<221> CDS

<222> (53)...(1393)

<223> codon_start: 53;
function: initiates complex N-glycans on secretory glycoproteins;
EC_number: 2.4.1.101;

<220>

<221> CDS

<222> (53)...(1393)

<223> product: beta-1,2-N-acetylglucosaminyltransferase I;

evidence: EXPERIMENTAL;

```
<220>
<221> CDS
<222> (53)...(1393)
<223> gene: cgl;
      standard_name: gntI;
      label: ORF;
      note: first gntI sequence from potato (unpublished).
```

```
<220>
<221> 5'UTR
<222> (15) ... (52)
```

```
<220>
<221> 3'UTR
<222> (1394)...(1655)
```

```
<220>
<221> CDS
<222> (80)...(139)
<223> function: membrane anchor (amino acids 10-29);
product: hydrophobic amino acid stretch in GnTI;
standard_name: membrane anchor of a type II Golgi
protein;
```

```
<220>
<221> CDS
<222> (80)...(139)
<223> note: identified by comparison with GnTI sequences
      from animals.
```

```
<220>
<221> misc_feature
<222> (1)...(14)
<223> function: used for cloning the cDNA library in
      Lambda ZAPII;
      product: EcoRI/NotI-cDNA adapter;
      number: 1.
```

```
<220>
<221> misc_feature
<222> (1656)...(1669)
<223> product: EcoRI/NotI-cDNA adapter;
      number: 2.
```

```
<400> 1
gaattcgcgg ccgcctgaga aaccctcgaa ttcaatttcg catttggcag ag atg aga 58
                                     Met Arg
                                     1
```

ggg aac aag ttt tgc ttt gat tta cgg tac ctt ctc gtc gtg gct gct 106
Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val Ala Ala

5	10	15	
ctc gcc ttc atc tac ata cag atg cgg ctt ttc gcg aca cag tca gaa			154
Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln Ser Glu			
20	25	30	
tat gta gac cgc ctt gct gct gca att gaa gca gaa aat cat tgt aca			202
Tyr Val Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His Cys Thr			
35	40	45	50
agt cag acc aga ttg ctt att gac aag att agc cag cag caa gga aga			250
Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln Gly Arg			
55	60	65	
gta gta gct ctt gaa gaa caa atg aag cat cag gac cag gag tgc cgg			298
Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu Cys Arg			
70	75	80	
caa tta agg gct ctt gtt cag gat ctt gaa agt aag ggc ata aaa aag			346
Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile Lys Lys			
85	90	95	
tta atc gga gat gtg cag atg cca gtg gca gct gta gtt gtt atg gct			394
Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val Met Ala			
100	105	110	
tgc agt cgt act gac tac ctg gag agg act att aaa tcc atc tta aaa			442
Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile Leu Lys			
115	120	125	130
tac caa aca tct gtt gca tca aaa tat cct ctt ttc ata tcc cag gat			490
Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln Asp			
135	140	145	
gga tca aat cct gat gta aga aag ctt gct ttg agc tat ggt cag ctg			538
Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly Gln Leu			
150	155	160	
acg tat atg cag cac ttg gat tat gaa cct gtg cat act gaa aga cca			586
Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu Arg Pro			
165	170	175	
ggg gaa ctg gtt gca tac tac aag att gca cgt cat tac aag tgg gca			634
Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp Ala			
180	185	190	
ttg gat cag ctg ttt cac aag cat aat ttt agc cgt gtt atc ata cta			682
Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile Ile Leu			
195	200	205	210
gaa gat gat atg gaa att gct gct gat ttt ttt gac tat ttt gag gct			730
Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe Glu Ala			
215	220	225	

gga gct act ctt ctt gac aga gac aag tcg att atg gct att tct tct 778
 Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile Ser Ser
 230 235 240

tgg aat gac aat gga caa agg cag ttc gtc caa gat cct gat gct ctt 826
 Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp Ala Leu
 245 250 255

tac cgc tca gac ttt ttt cct ggt ctt gga tgg atg ctt tca aaa tca 874
 Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser Lys Ser
 260 265 270

act tgg tcc gaa cta tct cca aag tgg cca aag gct tac tgg gat gac 922
 Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp Asp
 275 280 285 290

tgg cta agg ctg aaa gaa aat cac aga ggt cga caa ttt att cgc cca 970
 Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile Arg Pro
 295 300 305

gaa gtt tgc aga acg tac aat ttt ggt gag cat ggt tct agt ttg ggg 1018
 Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu Gly
 310 315 320

cag ttt ttt aag cag tat ctt gag cca att aag cta aat gat gtc cag 1066
 Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val Gln
 325 330 335

gtt gat tgg aag tca atg gac cta agt tac ctt ttg gag gac aac tat 1114
 Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp Asn Tyr
 340 345 350

gtg aaa cac ttt ggc gac ttg gtt aaa aag gct aag ccc atc cac gga 1162
 Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile His Gly
 355 360 365 370

gct gat gct gtt ttg aaa gca ttt aac ata gat ggt gat gtg cgt att 1210
 Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val Arg Ile
 375 380 385

cag tac aga gac caa cta gac ttt gaa gat atc gct cga cag ttt ggc 1258
 Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln Phe Gly
 390 395 400

att ttt gaa gaa tgg aag gat ggt gta cca cgg gca gca tat aaa ggg 1306
 Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr Lys Gly
 405 410 415

ata gta gtt ttc cgg ttt caa aca tct aga cgt gtg ttc ctt gtt tcc 1354
 Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu Val Ser
 420 425 430

cct gat tct ctt cga caa ctt gga gtt gaa gat act tag cgaagatatg 1403
 Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr *
 435 440 445

attggagcct gagcaacaat ttagacttat ttggtaggat acatttgaaa gagctgacac 1463
 gaaaagtatg actaccagta gctacatgca acattttaat gttaatggaa ggaacccact 1523
 gcttattgtt ggaatggatg aatcatcacc acatcctatt attcaagttt acaaacataa 1583
 agaggaaatg ttgccctata aaaacaaatt ttttgtttct aagaaggaac gttacgatta 1643
 tgagcaactt tggcggccgc gaattc 1669

<210> 2

<211> 446

<212> PRT

<213> Solanum tuberosum

<400> 2

Met Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val
 1 5 10 15
 Ala Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln
 20 25 30
 Ser Glu Tyr Val Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His
 35 40 45
 Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln
 50 55 60
 Gly Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu
 65 70 75 80
 Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
 85 90 95
 Lys Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val
 100 105 110
 Met Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile
 115 120 125
 Leu Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
 130 135 140
 Gln Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly
 145 150 155 160
 Gln Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu
 165 170 175
 Arg Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
 180 185 190
 Trp Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile
 195 200 205
 Ile Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe
 210 215 220
 Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
 225 230 235 240
 Ser Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp
 245 250 255
 Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
 260 265 270
 Lys Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
 275 280 285
 Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile

```

      290              295              300
Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
305              310              315              320
Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp
      325              330              335
Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
      340              345              350
Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
      355              360              365
His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
      370              375              380
Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln
385              390              395              400
Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
      405              410              415
Lys Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu
      420              425              430
Val Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr
      435              440              445

```

<210> 3
 <211> 1737
 <212> DNA
 <213> Nicotiana tabacum

<220>
 <221> misc_feature
 <222> (733)...(741)
 <223> function: Asn codon in this context is a potential
 glycosylation site;
 product: N-glycosylation consensus sequence;
 phenotype: N-glycans modulate protein properties;

<220>
 <221> misc_feature
 <222> (733)...(741)
 <223> standard_name: N-glycosylation site;
 label: pot-CHO;
 note: GnTI sequences from animals do not contain
 this feature.

<220>
 <221> CDS
 <222> (127)...(1467)
 <223> codon_start: 127;
 function: initiates complex N-glycans on secretory
 glycoproteins;
 EC_number: 2.4.1.101;

<220>
 <221> CDS

<222> (127)...(1467)
 <223> product: beta-1,2-N-acetylglucosaminyltransferase I;
 evidence: EXPERIMENTAL;
 gene: cgl;
 standard_name: gntI;

<220>

<221> CDS

<222> (127)...(1467)

<223> label: ORF;
 note: first gntI sequence from tobacco (unpublished).

<220>

<221> 5'UTR

<222> (15)...(126)

<220>

<221> 3'UTR

<222> (1468)...(1723)

<220>

<221> CDS

<222> (154)...(213)

<223> function: membrane anchor (amino acids 10-29);
 product: hydrophobic amino acid stretch in GntI;
 standard_name: membrane anchor of a type II golgi
 protein.

<220>

<221> misc_feature

<222> (1)...(14)

<223> function: use for cloning the cDNA library in
 Lambda ZAPII;
 product: EcoRI/NotI-cDNA adapter;
 number: 1.

<220>

<221> misc_feature

<222> (1724)...(1737)

<223> product: EcoRI/NotI-cDNA adapter;
 number: 2.

<400> 3

gaattcgcgg cgcgcattga cttgataccta actgaacagg caaagtaaata ccagcgatga 60
 aacactcata actgaacact gagagactat tcgctttctc ctaaagcctt caatcgaatt 120
 cgcacg atg aga ggg aac aag ttt tgc tgt gat ttc cgg tac ctc ctc 168
 Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu
 1 5 10

atc ttg gct gct gtc gcc ttc atc tac aca cag atg cgg ctt ttt gcg 216
 Ile Leu Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala
 15 20 25 30

aca cag tca gaa tat gca gat cgc ctt gct gct gca att gaa gca gaa 264
 Thr Gln Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu
 35 40 45

aat cat tgt aca agc cag acc aga ttg ctt att gac cag att agc ctg 312
 Asn His Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu
 50 55 60

cag caa gga aga ata gtt gct ctt gaa gaa caa atg aag cgt cag gac 360
 Gln Gln Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp
 65 70 75

cag gag tgc cga caa tta agg gct ctt gtt cag gat ctt gaa agt aag 408
 Gln Glu Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys
 80 85 90

ggc ata aaa aag ttg atc gga aat gta cag atg cca gtg gct gct gta 456
 Gly Ile Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val
 95 100 105 110

gtt gtt atg gct tgc aat cgg gct gat tac ctg gaa aag act att aaa 504
 Val Val Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys
 115 120 125

tcc atc tta aaa tac caa ata tct gtt gcg tca aaa tat cct ctt ttc 552
 Ser Ile Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe
 130 135 140

ata tcc cag gat gga tca cat cct gat gtc agg aag ctt gct ttg agc 600
 Ile Ser Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser
 145 150 155

tat gat cag ctg acg tat atg cag cac ttg gat ttt gaa cct gtg cat 648
 Tyr Asp Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His
 160 165 170

act gaa aga cca ggg gag ctg att gca tac tac aaa att gca cgt cat 696
 Thr Glu Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His
 175 180 185 190

tac aag tgg gca ttg gat cag ctg ttt tac aag cat aat ttt agc cgt 744
 Tyr Lys Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg
 195 200 205

gtt atc ata cta gaa gat gat atg gaa att gcc cct gat ttt ttt gac 792
 Val Ile Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp
 210 215 220

ttt ttt gag gct gga gct act ctt ctt gac aga gac aag tcg att atg 840
 Phe Phe Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met
 225 230 235

gct att tct tct tgg aat gac aat gga caa atg cag ttt gtc caa gat 888

Ala Ile Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp
 240 245 250

cct tat gct ctt tac cgc tca gat ttt ttt ccc ggt ctt gga tgg atg 936
 Pro Tyr Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met
 255 260 265 270

ctt tca aaa tct act tgg gac gaa tta tct cca aag tgg cca aag gct 984
 Leu Ser Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala
 275 280 285

tac tgg gac gac tgg cta aga ctc aaa gag aat cac aga ggt cga caa 1032
 Tyr Trp Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln
 290 295 300

ttt att cgc cca gaa gtt tgc aga aca tat aat ttt ggt gag cat ggt 1080
 Phe Ile Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly
 305 310 315

tct agt ttg ggg cag ttt ttc aag cag tat ctt gag cca att aaa cta 1128
 Ser Ser Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu
 320 325 330

aat gat gtc cag gtt gat tgg aag tca atg gac ctt agt tac ctt ttg 1176
 Asn Asp Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu
 335 340 345 350

gag gac aat tac gtg aaa cac ttt ggt gac ttg gtt aaa aag gct aag 1224
 Glu Asp Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys
 355 360 365

ccc atc cat gga gct gat gct gtc ttg aaa gca ttt aac ata gat ggt 1272
 Pro Ile His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly
 370 375 380

gat gtg cgt att cag tac aga gat caa cta gac ttt gaa aat atc gca 1320
 Asp Val Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala
 385 390 395

cgg caa ttt ggc att ttt gaa gaa tgg aag gat ggt gta cca cgt gca 1368
 Arg Gln Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala
 400 405 410

gca tat aaa gga ata gta gtt ttc cgg tac caa acg tcc aga cgt gta 1416
 Ala Tyr Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val
 415 420 425 430

ttc ctt gtt ggc cat gat tcg ctt caa caa ctc gga att gaa gat act 1464
 Phe Leu Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr
 435 440 445

taa caaagatatg attgcaggag cccgggcaaa atttttgact tattgggtag 1517

*

gatgcatcga gctgacacta aaccatgatt ttaccagtta catacaacgt tttaatgtta 1577
 tacggaggag ctcactgttc tagtggtgaa gggatatcgg cttcttagta ttggatgaat 1637
 catcaacaca acctattatt ttaagtgttc agaacataaa gaggaaatgt agccctgtaa 1697
 agactatata tgggaccatc ataatcgcg cgcgaattc 1737

<210> 4

<211> 446

<212> PRT

<213> Nicotiana tabacum

<400> 4

Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu Ile Leu
 1 5 10 15
 Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala Thr Gln
 20 25 30
 Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His
 35 40 45
 Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu Gln Gln
 50 55 60
 Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp Gln Glu
 65 70 75 80
 Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
 85 90 95
 Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val Val Val
 100 105 110
 Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys Ser Ile
 115 120 125
 Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
 130 135 140
 Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Asp
 145 150 155 160
 Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His Thr Glu
 165 170 175
 Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
 180 185 190
 Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg Val Ile
 195 200 205
 Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Phe Phe
 210 215 220
 Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
 225 230 235 240
 Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp Pro Tyr
 245 250 255
 Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
 260 265 270
 Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
 275 280 285
 Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
 290 295 300
 Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
 305 310 315 320
 Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp

```

                325                330                335
Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
                340                345                350
Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
                355                360                365
His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
                370                375                380
Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala Arg Gln
385                390                395                400
Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
                405                410                415
Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val Phe Leu
                420                425                430
Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr
                435                440                445

```

<210> 5
 <211> 1854
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <221> misc_feature
 <222> (1185)...(1193)
 <223> function: Asn Codon is a potential glycosylation site;
 product: Consensus sequence for N-glycosylation;
 phenotype: N glycans modulate protein characteristics;
 standard name: N glycosylation site;

<220>
 <221> misc_feature
 <222> (1185)...(1193)
 <223> label: pot-CHO;
 note: absent in animal GnTI sequences.

<220>
 <221> CDS
 <222> (135)...(1469)
 <223> codon_start: 135;
 function: initiates complex N glycans on secretory glycoproteins;
 EC_number: 2.4.1.101;

<220>
 <221> CDS
 <222> (135)...(1469)
 <223> product: beta-1,2-N-acetyl glucosaminyl transferase I;
 evidence: EXPERIMENTAL;
 gene: cgl;
 standard_name: gntI;

<220>
 <221> CDS
 <222> (135)...(1469)
 <223> label: ORF;

note: first gntI sequence from Arabidopsis
(unpublished).

<220>

<221> 5'UTR

<222> (19)...(134)

<220>

<221> 3'UTR

<222> (1470)...(1848)

<220>

<221> CDS

<222> (157)...(215)

<223> function: membrane anchor (amino acids 8-27);
product: hydrophobic amino-acid region in GntI;
standard_name: membrane anchor of a Type II Golgi protein;
note: identified by comparison with animal GntI sequences.

<220>

<221> misc_feature

<222> (1)...(18)

<223> function: for preparation of a cDNA library in
Lambda ACT;
product: XhoI-cDNA-Adaptor;
number: 1.

<220>

<221> misc_feature

<222> (1849)...(1854)

<223> product: XhoI-cDNA-Adaptor;
number: 2.

<400> 5

ctcgaggcca cgaaggccac cggtttttgtt ataacgaacg acaccgtttc aaacaacttc 60
cttattagct agtccctcc cgggcgcaaa caccagaaga tccaccgctt ttgatctggt 120
tgtttgctcg cgat atg gcg agg atc tcg tgt gac ttg aga ttt ctt ctc 170
Met Ala Arg Ile Ser Cys Asp Leu Arg Phe Leu Leu
1 5 10

atc ccg gca gct ttc atg ttc atc tac atc cag atg agg ctt ttc cag 218
Ile Pro Ala Ala Phe Met Phe Ile Tyr Ile Gln Met Arg Leu Phe Gln
15 20 25

acg caa tca cag tat gca gat cgc ctc agt tcc gct atc gaa tct gag 266
Thr Gln Ser Gln Tyr Ala Asp Arg Leu Ser Ser Ala Ile Glu Ser Glu
30 35 40

aac cat tgc act agt caa atg cga ggc ctc ata gat gaa gtt agc atc 314
Asn His Cys Thr Ser Gln Met Arg Gly Leu Ile Asp Glu Val Ser Ile
45 50 55 60

aaa cag tcg cgg att gtt gcc ctc gaa gat atg aag aac cgc cag gac 362

Lys Gln Ser Arg Ile Val Ala Leu Glu Asp Met Lys Asn Arg Gln Asp
 65 70 75
 gaa gaa ctt gtg cag ctt aag gat cta atc cag acg ttt gaa aaa aaa 410
 Glu Glu Leu Val Gln Leu Lys Asp Leu Ile Gln Thr Phe Glu Lys Lys
 80 85 90
 gga ata gca aaa ctc act caa ggt gga cag atg cct gtg gct gct gta 458
 Gly Ile Ala Lys Leu Thr Gln Gly Gly Gln Met Pro Val Ala Ala Val
 95 100 105
 gtg gtt atg gcc tgc agt cgt gca gac tat ctt gaa agg act gtt aaa 506
 Val Val Met Ala Cys Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys
 110 115 120
 tca gtt tta aca tat caa act ccc gtt gct tca aaa tat cct cta ttt 554
 Ser Val Leu Thr Tyr Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe
 125 130 135 140
 ata tct cag gat gga tct gat caa gct gtc aag agc aag tca ttg agc 602
 Ile Ser Gln Asp Gly Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser
 145 150 155
 tat aat caa tta aca tat atg cag cac ttg gat ttt gaa cca gtg gtc 650
 Tyr Asn Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val Val
 160 165 170
 act gaa agg cct ggt gaa ctg act gcg tac tac aag att gca cgt cac 698
 Thr Glu Arg Pro Gly Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His
 175 180 185
 tac aag tgg gca ctg gac cag ttg ttt tac aaa cac aaa ttt agt cga 746
 Tyr Lys Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg
 190 195 200
 gtg att ata cta gaa gac gat atg gaa att gct cca gac ttc ttt gat 794
 Val Ile Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp
 205 210 215 220
 tac ttt gag gct gca gct agt ctc atg gat agg gat aaa acc att atg 842
 Tyr Phe Glu Ala Ala Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met
 225 230 235
 gct gct tca tca tgg aat gat aat gga cag aag cag ttt gtg cat gat 890
 Ala Ala Ser Ser Trp Asn Asp Asn Gly Gln Lys Gln Phe Val His Asp
 240 245 250
 ccc tat gcg cta tac cga tca gat ttt ttt cct ggc ctt ggg tgg atg 938
 Pro Tyr Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met
 255 260 265
 ctc aag aga tcg act tgg gat gag tta tca cca aag tgg cca aag gct 986
 Leu Lys Arg Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala

270	275	280	
tac tgg gat gat tgg ctg aga cta aag gaa aac cat aaa ggc cgc caa			1034
Tyr Trp Asp Asp Trp Leu Arg Leu Lys Glu Asn His Lys Gly Arg Gln			
285	290	295	300
ttc att gca ccg gaa gtc tgt aga aca tac aat ttt ggt gaa cat ggg			1082
Phe Ile Ala Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly			
	305	310	315
tct agt ttg gga cag ttt ttc agt cag tat ctg gaa cct ata aag cta			1130
Ser Ser Leu Gly Gln Phe Phe Ser Gln Tyr Leu Glu Pro Ile Lys Leu			
	320	325	330
aac gat gtg acg gtt gac tgg aaa gca aag gac ctg gga tac ctg aca			1178
Asn Asp Val Thr Val Asp Trp Lys Ala Lys Asp Leu Gly Tyr Leu Thr			
	335	340	345
gag gga aac tat acc aag tac ttt tct ggc tta gtg aga caa gca cga			1226
Glu Gly Asn Tyr Thr Lys Tyr Phe Ser Gly Leu Val Arg Gln Ala Arg			
	350	355	360
cca att caa ggt tct gac ctt gtc tta aag gct caa aac ata aag gat			1274
Pro Ile Gln Gly Ser Asp Leu Val Leu Lys Ala Gln Asn Ile Lys Asp			
	365	370	375
gat gat cgt atc cgg tat aaa gac caa gta gag ttt gaa cgc att gca			1322
Asp Asp Arg Ile Arg Tyr Lys Asp Gln Val Glu Phe Glu Arg Ile Ala			
	385	390	395
ggg gaa ttt ggt ata ttt gaa gaa tgg aag gat ggt gtg cca cga aca			1370
Gly Glu Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Thr			
	400	405	410
gca tat aaa gga gta gtg gtg ttt cga atc cag aca aca aga cgt gta			1418
Ala Tyr Lys Gly Val Val Val Phe Arg Ile Gln Thr Thr Arg Arg Val			
	415	420	425
ttc ctg gtt ggg cca gat tct gta atg cag ctt gga att cga aat tcc			1466
Phe Leu Val Gly Pro Asp Ser Val Met Gln Leu Gly Ile Arg Asn Ser			
	430	435	440
tga tgcaaaacat atgaaaggaa aagaagattt tggaccgcat gcagcctcct			1519
*			
tctagcagct gttaggttgt attgttattt atggatgagt ttgtagagcg gtgggggttaa			1579
ctttaacagc aaggaagctc tggtgaccag gctgattggc ttagaagtta tgggaacccc			1639
ttgaaagggt caggggttaa tatatttcag ttgttttatt agtgattatc ttgtgggttaa			1699
cttatacgaa tgcaaatcat tctatgcagt ttttcttcgt ccacttggt ttggcttctc			1759
tattgctagt gtacatatct cttcaaaccat gtactaaata atgcgtgttg cttcaaagaa			1819
gtaactttta ttaaaaaaaaa aaaaaaaaaac tcgag			1854

<210> 6
 <211> 444
 <212> PRT
 <213> Arabidopsis thaliana

<400> 6
 Met Ala Arg Ile Ser Cys Asp Leu Arg Phe Leu Leu Ile Pro Ala Ala
 1 5 10 15
 Phe Met Phe Ile Tyr Ile Gln Met Arg Leu Phe Gln Thr Gln Ser Gln
 20 25 30
 Tyr Ala Asp Arg Leu Ser Ser Ala Ile Glu Ser Glu Asn His Cys Thr
 35 40 45
 Ser Gln Met Arg Gly Leu Ile Asp Glu Val Ser Ile Lys Gln Ser Arg
 50 55 60
 Ile Val Ala Leu Glu Asp Met Lys Asn Arg Gln Asp Glu Glu Leu Val
 65 70 75 80
 Gln Leu Lys Asp Leu Ile Gln Thr Phe Glu Lys Lys Gly Ile Ala Lys
 85 90 95
 Leu Thr Gln Gly Gly Gln Met Pro Val Ala Ala Val Val Val Met Ala
 100 105 110
 Cys Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys Ser Val Leu Thr
 115 120 125
 Tyr Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln Asp
 130 135 140
 Gly Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser Tyr Asn Gln Leu
 145 150 155 160
 Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val Val Thr Glu Arg Pro
 165 170 175
 Gly Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp Ala
 180 185 190
 Leu Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg Val Ile Ile Leu
 195 200 205
 Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Tyr Phe Glu Ala
 210 215 220
 Ala Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met Ala Ala Ser Ser
 225 230 235 240
 Trp Asn Asp Asn Gly Gln Lys Gln Phe Val His Asp Pro Tyr Ala Leu
 245 250 255
 Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Lys Arg Ser
 260 265 270
 Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp Asp
 275 280 285
 Trp Leu Arg Leu Lys Glu Asn His Lys Gly Arg Gln Phe Ile Ala Pro
 290 295 300
 Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu Gly
 305 310 315 320
 Gln Phe Phe Ser Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val Thr
 325 330 335
 Val Asp Trp Lys Ala Lys Asp Leu Gly Tyr Leu Thr Glu Gly Asn Tyr
 340 345 350
 Thr Lys Tyr Phe Ser Gly Leu Val Arg Gln Ala Arg Pro Ile Gln Gly
 355 360 365
 Ser Asp Leu Val Leu Lys Ala Gln Asn Ile Lys Asp Asp Arg Ile

```

          370          375          380
Arg Tyr Lys Asp Gln Val Glu Phe Glu Arg Ile Ala Gly Glu Phe Gly
385          390          395          400
Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Thr Ala Tyr Lys Gly
          405          410          415
Val Val Val Phe Arg Ile Gln Thr Thr Arg Arg Val Phe Leu Val Gly
          420          425          430
Pro Asp Ser Val Met Gln Leu Gly Ile Arg Asn Ser
          435          440

```

<210> 7
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer for cloning plant genes

<220>
 <221> misc_feature
 <222> (6)...(12)
 <223> n at positions 6, 9 & 12 = inosine.

<400> 7
 tgygynwsng cntggmayga yaay

24

<210> 8
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer for cloning plant genes

<220>
 <221> misc_feature
 <222> (4)...(10)
 <223> n at positions 4, 7, & 10 = inosine.

<220>
 <221> misc_feature
 <222> 13
 <223> n = a, c, g, or t.

<400> 8
 ccancntrn ccngsraara artc

24

<210> 9
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Primer for cloning plant genes

<220>

<221> VARIANT

<222> 2

<223> Xaa = Arg or Met.

<220>

<221> VARIANT

<222> 9

<223> Xaa = Asp or Tyr.

<400> 9

Gln Xaa Gln Phe Val Gln Asp Pro Xaa Ala Leu Tyr Arg Ser
1 5 10

<210> 10

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for cloning plant genes

<400> 10

atcggaagc ttggatcccc agtggcrgct gtagttgtta tggcttgc

48

<210> 11

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for cloning plant genes

<400> 11

ggccccccct cgaggtcgac ggtatcg

27

<210> 12

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for cloning plant genes

<400> 12

gggcctctag actccagagc yactactctt ccttgctgct ggctaattctt g

51

<210> 13

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for cloning plant genes

<400> 13

catggatccc tcgagaagcg tcaggaccag gagtgccggc

40

<210> 14

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for cloning plant genes

<400> 14

atcccgggat ccgctacgta tcttcaactc caagttg

37